



CAPSTONE PROJECT – DATA SCIENCE

MADRID AND BARCELONA NEIGHBORHOODS COMPARISON

CLAUDIO PRAVIA SILVERA – 12TH JUNE 2020

NEIGHBORHOOD COMPARISON BETWEEN MADRID AND BARCELONA

INTRODUCTION

We will compare the two cities to discover the neighborhoods for similarities related to the category and number of venues that they have.

BUSINESS PROBLEM

We want to provide franchise operators with information that makes it easier to identify the probability of success when opening a franchise in these cities.

DATA ACQUISITION AND CLEANSING

We get the list of neighborhoods of the two cities from Wikipedia, then we use *Nominatim* to obtain the geographical coordinates for the neighborhoods and finally we use Foursquare to obtain the different venues that exist in each of the neighborhoods.

DATA SOURCES

- The list of neighborhoods' data is obtained using the following Wikipedia links:
 - https://es.wikipedia.org/wiki/Anexo:Barrios_administrativos_de_Madrid
 - https://es.wikipedia.org/wiki/Distritos_de_Barcelona
- Nominatim
- Foursquare API

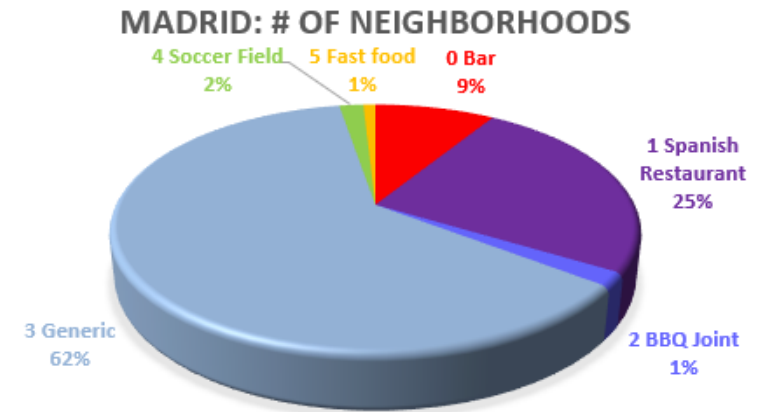
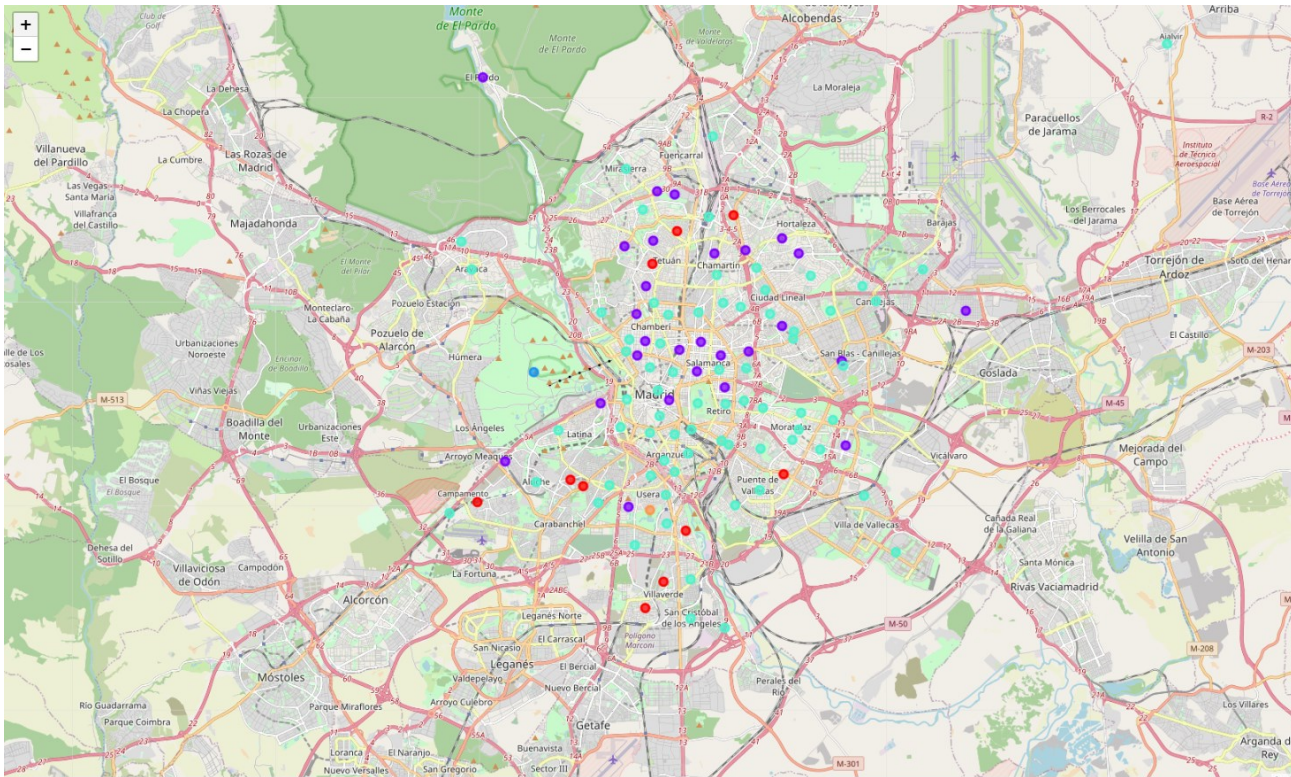
DATA CLEANSING

We discarded the neighborhoods for which we couldn't get the geographical coordinates.

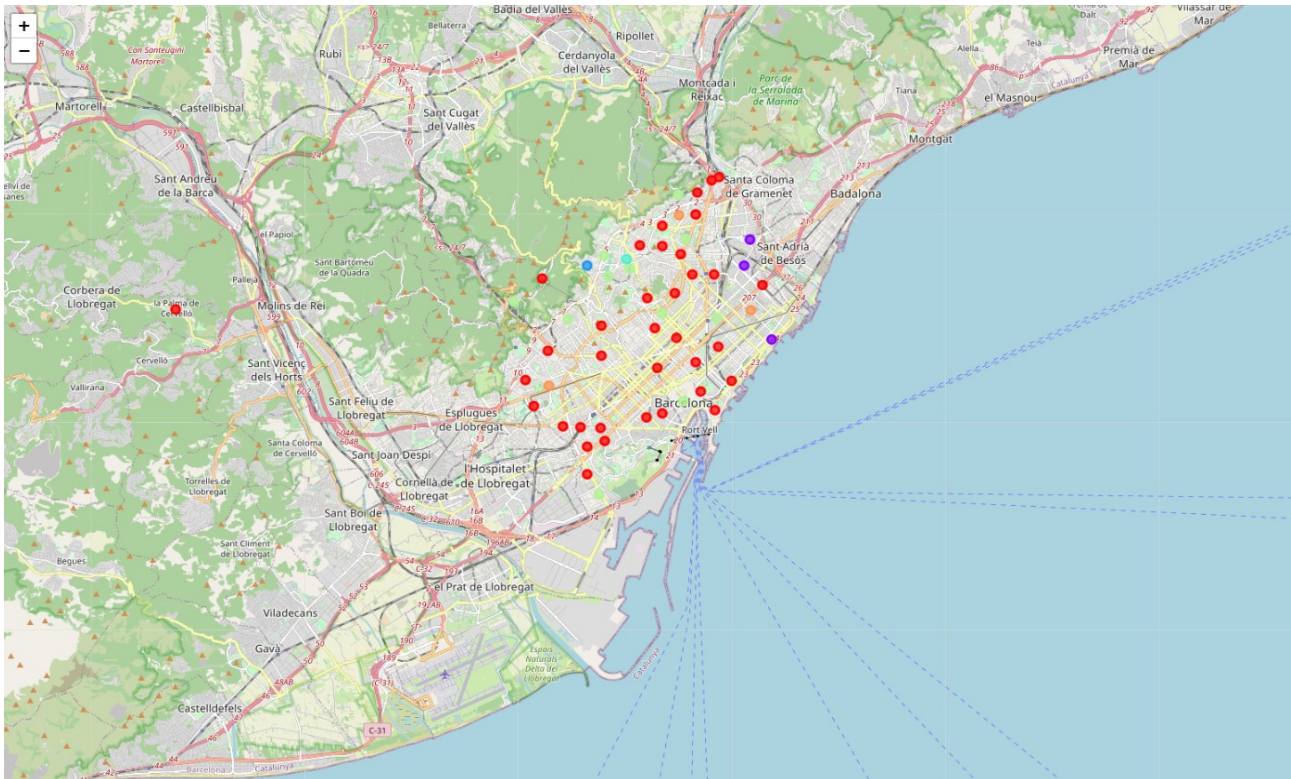
USING K-MEANS TO CLUSTER THE NEIGHBORHOODS

- Using Foursquare data, we identified the top 50 venues for each neighborhood that are within a 200-meter radius from the center of the neighborhood. We encoded and normalized the venues per neighborhood.
- Then, we got the 5 most common venues for each neighborhood.
- We used K-Means to cluster the neighborhoods in different groups to find patterns.
- This helped us get the information to solve our business problem.

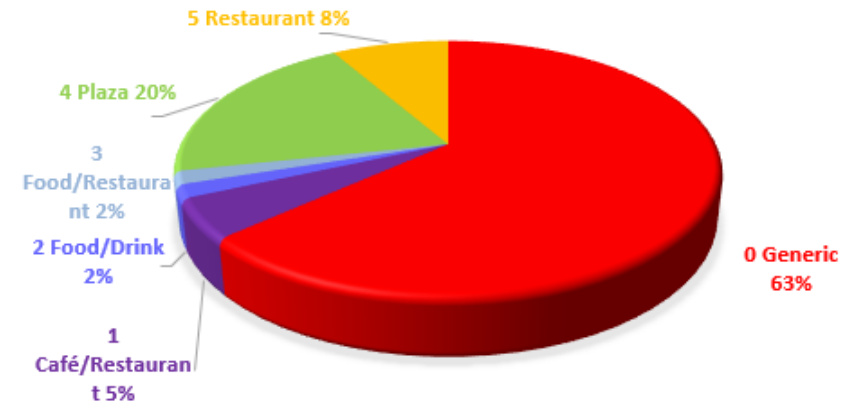
MADRID NEIGHBORHOODS



BARCELONA NEIGHBORHOODS



BARCELONA: # OF NEIGHBORHOODS



DISCUSSION

- In both cases there is a predominant cluster of more than 60% that contains venues that we labeled as Generic.
- In Madrid, the second largest cluster has a clear bias for the category “Spanish Restaurant” while the third has the category “Bar”. The “Spanish Restaurant” cluster is composed of the neighborhoods: Cortes, Ibiza, Recoletos, Lista, Castellana, Nueva España, Bellas Vistas, Valdeacederas, Arapiles, Almagro, Vallehermoso, El Pardo, La Paz, El Goloso, Valdezarza , Puerta del Ángel, Campamento, Comillas, Zofío, Ventas, Concepción, Atalaya, Canillas, Pinar del Rey, Apóstol Santiago, Valdebernardo, Simancas and Rejas.
- In the case of Barcelona, the second largest cluster has the category “Plaza” (meaning “park”) as predominant while the third has the category “Restaurant”. The “Restaurant” cluster is composed of the neighborhoods: Les Corts, La Salut, Canyelles, Verdún and Provençals del Poblenou

CONCLUSION AND FUTURE DIRECTIONS

- We can recommend, for a franchise in Spanish food that would like to open in Madrid, choose one of the neighborhoods in the Madrid cluster 1.
- In Barcelona, that is in Catalonia, a province that wants to be independent, there is no cluster related to Spanish Restaurant. The Barcelona cluster 5 though, shows neighborhoods with the main category as Restaurants. We can therefore recommend a more general food franchise to be opened in these neighborhoods.
- We could deepen this study by concentrating only in restaurants and then try to find more specific relationships between the type of cuisine and the neighborhoods in Madrid and Barcelona.



THANK YOU